D. ELLSBERG

ECON. DIV.

PSYCHOLOGICAL PROBABILITIES

L. J. Savage

P-95A

31 August 1949

The RAND Corporation
• SANTA MONICA • CALIFORNIA

RAND Paper P- 95-A

August 31, 1949

THE RAND CORPORATION
Santa Monica, California

Abstract of

PSYCHOLOGICAL PROBABILITIES

L. J. Savage

The notion that in some contexts the probability of an event should be understood to measure someone's personal mental degree of certainty that the event will take place is constantly recurring in discussions of the foundations of the theory of probability. This paper is an attempt to explore this notion in the light of the von Neumann-Morgenstern theory of utility. The author hopes that this work will have some value for the foundations of statistical inference and the theory of decision making.

To be presented at the meeting of the Econometric Society American Summer Session, Boulder Colorado, August 31, 1949

PSYCHOLOGICAL PROBABILITIES

1. Introduction

    1. Apology.

        1. No big new ideas and few little ones.

        2. Dusting off ideas which seem to have fallen into undeserved disuse.

    2. Historical background.

        1. With some dissent from the von Mises school the mathematical structure of probability is today generally agreed to be essentially that set forth by Kolmogoroff.

        2. Widespread disagreement about interpretation.

            1. Statisticians generally hold frequency interpretation.

            2. Certain scholars believe in necessary, and closely related logical interpretations. Keynes, Jeffries.

            3. Others in a personal or psychological interpretation. Bruno de Finetti and B. O. Koopman. The former's exposition especially attractive.

            4. Still others in an interplay of the foregoing interpretations. e.g. Carnap – frequency and logical.

            5. Not personally well acquainted with these controversies. Reviewed in Nagel's book.

            6. In this paper though actually in a state of doubt I shall for the sake of starting point take the frequency point of view as the only interpretation of probability admitting that what I am about to call psychological probability is not probability at all. This will I think suit most of my

audience, and I trust that others will not find the decision crucial to my actual thesis.

3. Central problem of modern statistical school as represented by Fisher, Neyman and Pearson, and Wald, has been to derive principles about the conduction of, interpretation of, and action based on experiments.
    1. This effort leaves one acquainted with it the impression that while a great deal has been accomplished there is at the same time a very serious hole unmendable by the usual techniques. Kendall's recent paper may be cited.
    2. The acme of this effort to date is in my opinion Wald's theory of minimum regret, and I shall accordingly take it as typifying the modern statistical school in the discussion to follow.

2. Utility:
    1. We shall need the von Neumann-Morgenstern theory of utility, which for the purpose at hand may be described thus.
        1. Let there be a class of objects X: x, y, z.
            1. x might be ten dollars or a fine kettle of fish, in short the traditional basket of goods.
            2. If a consumer is faced with choices among such "lottery tickets" as: $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$, the theory asserts that for him there is a numerical function u(x) such that of several lottery tickets he will prefer whichever **maximizes** $E(u(x))$, i.e., $\alpha_1 u(x) + \alpha_n u(x_n)$.
            3. This result is derived from assumptions about as palatable

as those of the traditional indifference curve analysis.

4. First it is to be emphasized that the theory really is a theory and that in any context it must stand or fall on its experimental test.

5. Second, as von Neumann frequently insists, the psychological assumptions on which the theory is based are quite naive, so that in many contexts we must expect it to fail. Nevertheless, for today's discussion I propose to accept the theory at least in certain contexts.

3. What is psychological probability?

   1. Language full of idioms expressing the feeling that not all uncertainties are of the same degree.

   2. We could rank the uncertainties of a given subject simply by asking him whether he considers A more likely than B. (Finetti)

      1. This is an objective procedure, but I don't like it because I don't believe that what a man says in response to a question which is not operationally meaningful to him will serve the purpose I have in mind.

         (i.e., Savage asks: How does he arrive at his answer, or check it, or "correct" it if inconsistencies arise?)

      2. Let us rather say to the subject:
         "You may have a dollar if A proves true or else a dollar if B proves true. Which ~~do~~ would you prefer?"

         (Savage is "really interested" in how he would bet, anyway; he is interested in man's beliefs only insofar as they would affect his bets; so ask him about bets; or watch him bet.)

   3. Let us examine this program in more detail.

      1. X might be the class of all envisaged future histories of the "world".

         1. Of course in some contexts the "world" is very simple.

         2. Events are then subsets of X.

3. The test of $A \leq B$

"A not more likely than B" is to determine whether the subject would prefer a particular gift if A materializes or the same gift if B materializes.

1. Introspection suggests that as long as the gift represents some improvement over the subject's status quo it scarcely matters what or how big it is, e.g., prestige, or a lottery ticket.
2. If an experiment of this sort were conducted some care would have to be taken that the gift be something worth thinking for. The subject's responses would be less influenced by random variation if the gift is large.

4. The following axioms are put forth as plausible.

1. " $\leq$ " is a simple ordering.
2. $0 \leq A$ for all A.
3. If $A \leq B$ and $A \cap C = B \cap C = 0$, then $A \cup C \leq B \cup C$, and conversely.
4. There are arbitrarily large integers n for which there exist partitions $A_1, \ldots, A_n$ with $A_i \simeq A_n$.

5. Criticisms of these axioms.

1. Axiom 1 is unrealistic insofar as it does not take proper account of vagueness and "mistakes", i.e., breaches of the axiom, which the subject himself would revise if called to his attention.
2. Axiom 2 seems very solid to me.
3. Axiom 3 also looks good.
4. Axiom 4 would in practice be realised, if for example the subject knew he had a fair coin in the frequency sense.

6. Consequences of the axioms.

   1. I gather from de Finetti, though he alludes to a fifth axiom which does not seem relevant just here, that in the presence of these axioms the definition of numerical psychological probability suggested by them is consistent and is a finitely additive measure on all the subsets of X.

   2. Technical note: Unless the discussion be confined to a a reasonable algebra of the subsets of X, this general result, though true, might prove annoying. For example a subject who wished to assign equal psychological probability to congruent sets or the surface of a sphere would be quite annoyed at Hausdorff.

4. Direct measurement of probability.

   1. The idea of measuring psychological probability by investigating what bets the subject is willing to make is old. de Finetti, for example traces it to Bertrand.

   2. If utility were linear in money the task would be straight forward, and is indeed carried out in detail by Finetti: de Finetti's main technical device here is the requirement that the bets which the subject is willing to make should not be compoundable into a sure loss. Coherence

      No; still problem of ambiguity

   3. If the utility function of a subject were measured the task would again be essentially the same.

   4. Again if small bets could be used it would be easy.

   5. But small bets, like small gifts, would not provide a good experimental technique, and utility functions are very difficult

to measure, so this method seems to compound difficulties unnecessarily. However I do not take this to be the last word on the subject.

4. The consistency requirements.

   1. Introspection shows that the theory of psychological probability will not be strictly true for any human subject.

      For example though I attach the same psychological probability to every permutation of a deck of cards, it is not likely that I attach to the psychological probability of a full house anything like the only value which is consistent with the first evaluation.

   2. As a matter of fact we revise our own evaluations in the light of the axioms and regard it as a "mistake" to be in disagreement with them. Thus the axioms are not only a theory of behavior, but are by most of us interpreted as normative.

   3. An analogy with logic is often pointed out in that axioms of logic may be taken as a prediction of relationship among the propositions which a subject on interrogation will be found to believe (without doubt), and also as a normative code.

5. The decision problem.

   1. Very generally a decision may be regarded as a choice of one out of several functions $a_i(x)$ defined over the envisaged future histories of the world and taking as values "baskets of goods".

      1. Much as in the von Neumann theory of games it must be recognized that the choice of i may imply the choice of an actually elaborate policy.

      2. We are concerned both with theories about how such decisions will be made by an individual, and about how they should be made.

6. Solution by psychological probability.

   1. Utility can be derived for psychological probabilities just as for real probabilities, so that $a_i(x)$ can be assumed to take utility values.

   2. Then the combined psychological theories of utility and psychological probability predict that the subject will choose that i for which the expected value of $a_i(x)$ is maximized.

   3. At the same time the theory can be interpreted normatively.

7. Wald's maxim.

   1. The definition

      1. Replace the values of $a_i(x)$ by their utilities based on real probability.

      2. Let

         $$r_i(x) = \max_j a_j(x) - a_i(x)$$

         1. I believe that Wald himself does not attach great importance to this step, though it seems very important to me.

      3. Choose a distribution of i such that

         $$\max_x \sum r_i(x) p(i)$$

         is as small as possible.

   2. This theory also is both predictive and normative.

   3. The practical motivation.

      1. Such a player makes "nearly" as much as can be made.

8. Comparison of two solutions.

   1. Psychological probability precludes the use of randomization as a tool which seems absurd.

2. The minimum risk principle leads to absurd results wherever the the subject is very sure of a highly relevant fact.

   1. Example. If offered either side of an even money bet that the German government will become a monarchy in the next five years, the theory of minimum risk tells me to decide by the flip of a coin.

3. One very attractive feature of the minimum risk theory is its public character.

4. For the time, it seems to me that both theories are being and should be used side by side.

LJS:je